

AMENDMENTS TO THE CLAIMS

This listing of claims will replace all prior versions, and listings, of claims in the application.

Claim underlining shows the changes from the originally issued patent.

1-26. (Cancelled)

27. (Thrice Amended) A method of parallelizing an operation, the method comprising the steps of:
dividing the operation into a set of work partitions;
assigning work partitions from said set of work partitions to a plurality of entities,
wherein at least one entity of said plurality of entities is assigned a plurality of
work partitions from said set of work partitions;
wherein the step of assigning work partitions is performed by assigning the work
partitions in a sequence based at least in part on sizes associated with the work
partitions;
said plurality of entities operating in parallel on work partitions assigned to them to
perform said operation; and
wherein assigning the work partitions in a sequence includes assigning a first previously
unassigned work partition to a particular entity of the plurality of entities, and
when the particular entity completes processing the first work partition, picking a
second previously unassigned work partition based at least in part to the size of
the second work partition, and assigning the second unassigned work partition to
the particular entity for processing.

28. (Once Amended) The method of Claim 27 wherein the step of assigning the work
partitions in a sequence is performed by assigning relatively larger work partitions before
assigning relatively smaller work partitions.

29. (Thrice Amended) A method of parallelizing an operation, the method comprising the steps of:
dividing the operation into a set of work partitions;
assigning work partitions from said set of work partitions to a plurality of entities,
wherein at least one entity of said plurality of entities is assigned a plurality of
work partitions from said set of work partitions, wherein the step of assigning
work partitions includes:
assigning said at least one entity a first work partition from said set of work
partitions; and
after said at least one entity has completed operation on said first work partition,
assigning said at least one entity a second work partition from said set of work
partitions, wherein the step of assigning said at least one entity a second work
partition includes
determining whether there are any unassigned work partitions from a first level in
a hierarchy to which said first work partition belonged; and
if there are no unassigned work partitions from the first level in the
hierarchy, then selecting said second work partition from a level in
said hierarchy that is two levels above said first level in said
hierarchy;
said plurality of entities operating in parallel on work partitions assigned to them to
perform said operation; and
wherein the operation is specified in a query that corresponds to the hierarchy of
operations.

30. (Once Amended) A method of parallelizing an operation, the method comprising the steps of:
dividing the operation into a set of work partitions;
assigning work partitions from said set of work partitions to a plurality of entities,
wherein at least one entity of said plurality of entities is assigned a plurality of
work partitions from said set of work partitions;

7 said plurality of entities operating in parallel on work partitions assigned to them to
 8 perform said operation;
 9 the method includes the step of generating a serial execution plan for operations in a
 10 database management system (DBMS) running on a computer system;
 11 the method includes the step of generating a parallelized execution plan for said serial
 12 execution plan, said parallelized execution plan including first and second
 13 operations;
 14 the step of dividing an operation is performed by dividing said second operation;
 15 the plurality of entities includes one or more slave processes operating on a plurality of
 16 data partitions, the quantity of said data partitions being greater than the quantity
 17 of said slave processes;
 18 executing said parallelized execution plan when a plurality of parallel resources of said
 19 computer system are available; and
 20 executing said serial execution plan when said plurality of resources are not available.

1 31. (Once Amended) The method of claim 30 wherein said step of generating a parallelized
 2 execution plan includes the steps of:
 3 identifying one or more segments of said serial execution plan that can be parallelized;
 4 and
 5 identifying partitioning requirements of said one or more segments.

1 32. (Once Amended) The method of claim 30 wherein said step of generating a parallelized
 2 execution plan is based on a specification of parallelism in a statement specifying one of
 3 said operations.

1 33. (Once Amended) A method of parallelizing an operation, the method comprising the
 2 steps of:
 3 dividing the operation into a set of work partitions;

assigning work partitions from said set of work partitions to a plurality of entities,
wherein at least one entity of said plurality of entities is assigned a plurality of
work partitions from said set of work partitions;
said plurality of entities operating in parallel on work partitions assigned to them to
perform said operation;
generating an execution plan for said operation;
examining said execution plan from bottom up;
identifying a parallelized portion of said execution plan, said parallelized portion can be
processed in parallel, said parallelized portion including first and second
operations, said first and second operations being executable in parallel;
wherein the step of dividing the operation is performed by dividing said second operation;
wherein the plurality of entities includes one or more slave processes operating on a
plurality of data partitions, the quantity of said data partitions being greater than
the quantity of said slave processes;
identifying some serial portion of said execution plan, said serial portion can be processed
in serial; and
allocating a central scheduler between said parallelized portion and said serial portion.

34. (Once Amended) The method of Claim 33 further including the steps of:
identifying a first data flow requirement for a first portion of said execution plan said first
data flow requirement corresponding to a partitioning of a data flow required by
said first portion;
identifying a second data flow requirement for a second portion of said execution plan
said second data flow requirement corresponding by said second portion; and
allocating a data flow director between said first portion and said second portion when
said first data flow requirement is not compatible with said second data flow
requirement said data flow director repartitioning a data flow of said first portion
to be compatible with said second data flow requirement.

35. (Once Amended) A method for parallelizing an operation, the method comprising the steps of:
dividing the operation into a set of work partitions;
assigning work partitions from said set of work partitions to a plurality of entities,
wherein at least one entity of said plurality of entities is assigned a plurality of
work partitions from said set of work partitions;
said plurality of entities operating in parallel on work partitions assigned to them to
perform said operation;
generating an execution plan to execute database management system (DBMS) operations
in parallel, said execution plan including first and second operations;
wherein the step of dividing said operation is performed by dividing said second
operation;
initiating an operation coordinator in a computer system to coordinate execution of said
execution plan;
initiating, by said operation coordinator, a first set of slaves operating on a plurality of
data partitions to produce data, the quantity of said data partitions being greater
than the quantity of said first set of slave processes;
initiating, as said plurality of entities, by said operation coordinator, a second set of slaves
to consume data; and
directing said second set of slaves to produce data and said first set of slaves to consume
data when said first set of slaves finishes producing data.

36. (Once Amended) The method of claim 35 wherein said execution plan is comprised of operator nodes and said operator nodes are linked together to form execution sets.

37. (Once Amended) A method for parallelizing an operation, the method comprising the steps of:
dividing the operation into a set of work partitions;

assigning work partitions from said set of work partitions to a plurality of entities,
wherein at least one entity of said plurality of entities is assigned a plurality of
work partitions from said set of work partitions;
said plurality of entities operating in parallel on work partitions assigned to them to
perform said operation;
generating an execution plan to execute said operations in parallel, said execution plan
including first and second operations;
wherein the step of dividing said operation includes dividing said first operation;
initiating producer slaves operating on a plurality of data partitions to produce a first data
production;
initiating consumer slaves to consume said first data production;
when said first data production is completed, generating an identification of a plurality of
said consumer slaves that did not receive data in said first data production;
examining said identification during a subsequent data production; and
reducing said subsequent data production such that said subsequent data production does
not produce data for said plurality of said consumer slaves.

38. (Once Amended) A method for processing a query, the method comprising the steps of:
receiving a statement that specifies at least an operation;
determining a user-specified degree of parallelism to use in performing the operation;
dividing the operation into a set of work partitions;
performing a determination of how many entities to use to perform said operation based,
at least in part, on the user-specified degree of parallelism;
assigning work partitions from said set of work partitions to a plurality of entities based
on said determination; and
said plurality of entities operating in parallel on work partitions assigned to them to
perform said operation.

39. (Once Amended) The method of Claim 38 wherein:
the query requires a plurality of operations;

the user-specified degree of parallelism is specified in said statement, and
the statement specifies said degree of parallelism for a subset of the plurality of
operations required by the query.

40. (Once Amended) The method of Claim 39 wherein
the user-specified degree of parallelism is specified in said statement; and
the degree of parallelism specified by the query indicates that no amount of parallelism is
to be used during execution of a particular portion of the query.

41. (Once Amended) The method of Claim 38 wherein
the user-specified degree of parallelism is specified in said statement, and
the degree of parallelism specified by the query indicates a maximum amount of
parallelism to use during execution of said operation.

42. (Once Amended) A method of processing a query, the method comprising the steps of:
dividing an operation required by said query into a set of work partitions by generating a
set of query fragments;
incorporating hints into at least some of said query fragments, wherein the hint associated
with a given query fragment indicates how to perform the work partition
associated with said given query fragment;
assigning query fragments from said set of query fragments to a plurality of entities; and
said plurality of entities operating in parallel on query fragments assigned to them to
perform said operation, wherein entities working on a query fragment associated
with a hint perform the work partition associated with said query fragment in a
manner dictated by said hint.

43. (Once Amended) The method of Claim 42 wherein the step of incorporating hints
includes incorporating hints that dictate the operation of a table scan.

44. (Once Amended) The method of Claim 43 wherein the step of incorporating hints that dictate the operation of a table scan includes incorporating hints that rowid partitioning is to be used during the table scan.

45. (Once Amended) The method of Claim 42 wherein the step of incorporating hints includes incorporating hints that specify performance of a full table scan.

46. (Once Amended) The method of Claim 42 wherein the step of incorporating hints includes incorporating hints that specify using a particular type of join.

47. (Once Amended) The method of Claim 46 wherein the step of incorporating hints that specify using a particular type of join includes incorporating hints that specify using a sort/merge join.

48. (Once Amended) The method of Claim 46 wherein the step of incorporating hints that specify using a particular type of join includes incorporating hints that specify using a nested loop join.

49. (Once Amended) A method of processing a query, the method comprising the steps of:
determining a hierarchy of operations associated with a query;
dividing a first operation required by said query into a first set of work partitions;
dividing a second operation required by said query into a second set of work partitions,
wherein said second operation immediately follows said first operation in said
hierarchy;
dividing a third operation required by said query into a third set of work partitions,
wherein said third operation immediately follows said second operation in said
hierarchy;
assigning work partitions from said first set of work partitions to a first plurality of
entities;

said first plurality of entities operating in parallel on work partitions assigned to them
from said first set of work partitions to perform said first operation;
assigning work partitions from said second set of work partitions to a second plurality of
entities, wherein said second plurality of entities are different entities than said
first plurality of entities; and
said second plurality of entities operating in parallel on work partitions assigned to them
from said second set of work partitions to perform said second operation;
assigning work partitions from said third set of work partitions to said first plurality of
entities; and
said first plurality of entities operating in parallel on work partitions assigned to them
from said third set of work partitions to perform said third operation.

50. (Once Amended) The method of Claim 49 further comprising performing the following
steps when a given entity in said first set of entities finishes performing a work partition
from said first set of work partitions:
determining whether there are any unassigned work partitions from said first set of work
partitions; and
if there are no unassigned work partitions from said first set of work partitions, then
assigning the given entity a work partition selected from said third set of work
partitions; and
if there are unassigned work partitions from said first set of work partitions, then
assigning the given entity a work partition selected from said first set of work
partitions.

51. (Once Amended) The method of Claim 49 wherein the hierarchy includes odd levels and
even levels, and the method further comprises the steps of assigning work partitions from
odd levels to said first plurality of entities and work partitions from even levels to said
second plurality of entities.

52. (Once Amended) The method of Claim 49 wherein performing work partitions in said first set of work partitions causes said first set of entities produce output consumed by said second plurality of entities, and performing work partitions in said third set of work partitions causes said first set of entities to consume output produced by said second plurality of entities.

53-62. (Cancelled)

63. (Twice Amended) A computer-readable storage medium carrying instructions for parallelizing an operation, the instructions including instructions for performing the steps of:
dividing the operation into a set of work partitions;
assigning work partitions from said set of work partitions to a plurality of entities,
wherein at least one entity of said plurality of entities is assigned a plurality of work partitions from said set of work partitions;
wherein the step of assigning work partitions is performed by assigning the work partitions in a sequence based at least in part on sizes associated with the work partitions;
said plurality of entities operating in parallel on work partitions assigned to them to perform said operation; and
wherein assigning the work partitions in a sequence includes assigning a first previously unassigned work partition to a particular entity of the plurality of entities, and when the particular entity completes processing the first work partition, picking a second previously unassigned work partition based at least in part to the size of the second work partition, and assigning the second unassigned work partition to the particular entity for processing.

64. (Twice Amended) The computer-readable storage medium of Claim 63 wherein the step of assigning the work partitions in a sequence is performed by assigning relatively larger work partitions before assigning relatively smaller work partitions.

65. (Twice Amended) A computer-readable storage medium carrying instructions for parallelizing an operation, the instructions including instructions for performing the steps of:
dividing the operation into a set of work partitions;
assigning work partitions from said set of work partitions to a plurality of entities,
wherein at least one entity of said plurality of entities is assigned a plurality of work partitions from said set of work partitions, wherein the step of assigning work partitions includes
assigning said at least one entity a first work partition from said set of work partitions;
and
after said at least one entity has completed operating on said first work partition,
assigning said at least one entity a second work partition from said set of work partitions;
said plurality of entities operating in parallel on work partitions assigned to them to perform said operation;
wherein the operation is specified in a query that corresponds to a hierarchy of operations;
and
the step of assigning said at least one entity a second work partition includes
determining whether there are any unassigned work partitions from a first level in the hierarchy to which said first work partition belonged; and
if there are no unassigned work partitions from the first level in the hierarchy, then
selecting said second work partition from a level in said hierarchy that is
two levels above said first level in said hierarchy.

66. (Twice Amended) A computer-readable storage medium carrying instructions for parallelizing an operation, the instructions including instructions for performing the steps of:
dividing the operation into a set of work partitions;

5 assigning work partitions from said set of work partitions to a plurality of entities,
 6 wherein at least one entity of said plurality of entities is assigned a plurality of
 7 work partitions from said set of work partitions;
 8 said plurality of entities operation in parallel on work partitions assigned to them to
 9 perform said operation;
 10 wherein the instructions include instructions for performing the step of generating a serial
 11 execution plan for operations in a database management system (DBMS) running
 12 on a computer system;
 13 wherein the instructions include instructions for performing the step of generating a
 14 parallelized execution plan for said serial execution plan, said parallelized
 15 execution plan including first and second operations;
 16 wherein the step of dividing an operation is performed by dividing said second operation;
 17 wherein the plurality of entities includes one or more slave processes operating on a
 18 plurality of data partitions, the quantity of said data partitions being greater than
 19 the quantity of said slave processes;
 20 wherein the instructions include instructions for performing the step of executing said
 21 parallelized execution plan when a plurality of parallel resources of said computer
 22 system are available; and
 23 wherein the instructions include instructions for performing the step of executing said
 24 serial execution plan when said plurality of resources are not available.

1 67. (Twice Amended) The computer-readable storage medium of claim 66 wherein said step
 2 of generating a parallelized execution plan includes the steps of:
 3 identifying one or more segments of said serial execution plan that can be parallelized;
 4 and
 5 identifying partitioning requirements of said one or more segments.

1 68. (Twice Amended) The computer-readable storage medium of claim 66 wherein said step
 2 of generating a parallelized execution plan is based on a specification of parallelism in a
 3 statement specifying one of said operations.

69. (Twice Amended) A computer-readable storage medium carrying instructions for parallelizing an operation, the instructions including instructions for performing the steps of:
dividing the operation into a set of work partitions;
assigning work partitions from said set of work partitions to a plurality of entities,
wherein at least one entity of said plurality of entities is assigned a plurality of
work partitions from said set of work partitions;
said plurality of entities operating in parallel on work partitions assigned to them to
perform some operation;
generating an execution plan for said operation;
examining said execution plan from bottom up;
identifying a parallelized portion of said execution plan, said parallelized portion can be
processed in parallel, said parallelized portion including first and second
operations, said first and second operations being executable in parallel;
wherein the step of dividing the operation is performed by dividing said second operation;
wherein the plurality of entities includes one or more slave processes operating on a
plurality of data partitions, the quantity of said data partitions being greater than
the quantity of said slave processes;
identifying some serial portion of said execution plan, said serial portion can be processed
in serial; and
allocating a central scheduler between said parallelized portion and said serial portion.

70. (Twice Amended) The computer-readable storage medium of Claim 69 further including instructions for performing the steps of:
identifying a first data flow requirement for a first portion of said execution plan said first
data flow requirement corresponding to a partitioning of a data flow required by
said first portion;
identifying a second data flow requirement for a second portion of said execution plan
said second data flow requirement corresponding by said second portion; and

allocating a data flow director between said first portion and said second portion when said first data flow requirement is not compatible with said second data flow requirement said data flow director repartitioning a data flow of said first portion to be compatible with said second data flow requirement.

71. (Twice Amended) A computer-readable storage medium carrying instructions for parallelizing an operation, the instructions including instructions for performing the steps of:
dividing the operation into a set of work partitions;
assigning work partitions from said set of work partitions to a plurality of entities,
wherein at least one entity of said plurality of entities is assigned a plurality of work partitions from said set of work partitions;
said plurality of entities operating in parallel on work partitions assigned to them to perform said operation;
generating an execution plan to execute database management system (DBMS) operations in parallel, said execution plan including first and second operations;
wherein the step of dividing said operation is performed by dividing said second operation;
initiating an operation coordinator in a computer system to coordinate execution of said execution plan;
initiating, by said operation coordinator, a first set of slaves operating on a plurality of data partitions to produce data, the quantity of said data partitions being greater than the quantity of said first set of slave processes;
initiating, as said plurality of entities, by said operation coordinator, a second set of slaves to consume data; and
directing said second set of slaves to produce data and said first set of slaves to consume data when said first set of slaves finishes producing data.

72. (Twice Amended) The computer-readable storage medium of claim 71 wherein said execution plan is comprised of operator nodes and said operator nodes are linked together to form execution sets.

73. (Twice Amended) A computer-readable storage medium carrying instructions for parallelizing an operation, the instructions including instructions for performing the steps of:
dividing the operation into a set of work partitions;
assigning work partitions from said set of work partitions to a plurality of entities,
wherein at least one entity of said plurality of entities is assigned a plurality of work partitions from said set of work partitions;
said plurality of entities operating in parallel on work partitions assigned to them to perform said operation;
generating an execution plan to execute said operations in parallel, said execution plan including first and second operations;
wherein the step of dividing said operation includes dividing said first operation;
initiating producer slaves operating on a plurality of data partitions to produce a first data production;
initiating consumer slaves to consume said first data production;
when said first data production is completed, generating an identification of a plurality of said consumer slaves that did not receive data in said first data production;
examining said identification during a subsequent data production; and
reducing said subsequent data production such that said subsequent data production does not produce data for said plurality of said consumer slaves.

74. (Twice Amended) A computer-readable storage medium storing instructions for processing a query, the instructions including instructions for performing the steps of:
receiving a statement that specifies at least an operation;
determining a user-specified degree of parallelism to use in performing the operation;
dividing the operation into a set of work partitions;

performing a determination of how many entities to use to perform said operation based,
at least in part, on the user-specified degree of parallelism;
assigning work partitions from said set of work partitions to a plurality of entities based
on said determination; and
said plurality of entities operating in parallel on work partitions assigned to them to
perform said operation.

75. (Twice Amended) The computer-readable storage medium of Claim 74 wherein:
the query requires a plurality of operations;
the user-specified degree of parallelism is specified in said statement, and
the statement specifies said degree of parallelism for a subset of the plurality of
operations required by the query.

76. (Twice Amended) The computer-readable storage medium of Claim 75 wherein
the user-specified degree of parallelism is specified in said statement; and
the degree of parallelism specified by the query indicates that no amount of parallelism is
to be used during execution of a particular portion of the query.

77. (Twice Amended) The computer-readable storage medium of Claim 74 wherein
the user-specified degree of parallelism is specified in said statement, and
the degree of parallelism specified by the query indicates a maximum amount of
parallelism to use during execution of said operation.

78. (Twice Amended) A computer-readable storage medium carrying instructions for
processing a query, the instructions including instructions for performing the steps of:
dividing an operation required by said query into a set of work partitions by generating a
set of query fragments;
incorporating hints into at least some of said query fragments, wherein the hint associated
with a given query fragment indicates how to perform the work partition
associated with said given query fragment;

assigning query fragments from said set of query fragments to a plurality of entities; and
said plurality of entities operating in parallel on query fragments assigned to them to
perform said operation, wherein entities working on a query fragment associated
with a hint perform the work partition associated with said query fragment in a
manner dictated by said hint.

79. (Twice Amended) The computer-readable storage medium of Claim 78 wherein the step
of incorporating hints includes incorporating hints that dictate the operation of a table
scan.

80. (Twice Amended) The computer-readable storage medium of Claim 79 wherein the step
of incorporating hints that dictate the operation of a table scan includes incorporating
hints that rowid partitioning is to be used during the table scan.

81. (Twice Amended) The computer-readable storage medium of Claim 78 wherein the step
of incorporating hints includes incorporating hints that specify performance of a full table
scan.

82. (Twice Amended) The computer-readable storage medium of Claim 78 wherein the step
of incorporating hints includes incorporating hints that specify using a particular type of
join.

83. (Twice Amended) The computer-readable storage medium of Claim 82 wherein the step
of incorporating hints that specify using a particular type of join includes incorporating
hints that specify using a sort/merge join.

84. (Twice Amended) The computer-readable storage medium of Claim 82 wherein the step
of incorporating hints that specify using a particular type of join includes incorporating
hints that specify using a nested loop join.

85. (Twice Amended) A computer-readable storage medium carrying instructions for processing a query, the instructions including instructions for performing the steps of: determining a hierarchy of operations associated with a query; dividing a first operation required by said query into a first set of work partitions; dividing a second operation required by said query into a second set of work partitions, wherein said second operation immediately follows said first operation in said hierarchy; dividing a third operation required by said query into a third set of work partitions, wherein said third operation immediately follows said second operation in said hierarchy; assigning work partitions from said first set of work partitions to a first plurality of entities; said first plurality of entities operating in parallel on work partitions assigned to them from said first set of work partitions to perform said first operation; assigning work partitions from said second set of work partitions to a second plurality of entities, wherein said second plurality of entities are different entities than said first plurality of entities; and said second plurality of entities operating in parallel on work partitions assigned to them from said second set of work partitions to perform said second operation; assigning work partitions from said third set of work partitions to said first plurality of entities; and said first plurality of entities operating in parallel on work partitions assigned to them from said third set of work partitions to perform said third operation.

86. (Twice Amended) The computer-readable storage medium of Claim 85 further comprising instructions for performing the following steps when a given entity in said first set of entities finishes performing a work partition from said first set of work partitions: determining whether there are any unassigned work partitions from said first set of work partitions; and

if there are no unassigned work partitions from said first set of work partitions, then
assigning the given entity a work partition selected from said third set of work
partitions; and
if there are unassigned work partitions from said first set of work partitions, then
assigning the given entity a work partition selected from said first set of work
partitions.

87. (Twice Amended) The computer-readable storage medium of Claim 85 wherein the
hierarchy includes odd levels and even levels, and the instructions further include
instructions for performing the steps of assigning work partitions from odd levels to said
first plurality of entities and work partitions from even levels to said second plurality of
entities.

88. (Twice Amended) The computer-readable storage medium of Claim 85 wherein
performing work partitions in said first set of work partitions causes said first set of
entities produce output consumed by said second plurality of entities, and performing
work partitions in said third set of work partitions causes said first set of entities to
consume output produced by said second plurality of entities.

89-91. (Canceled).

92. (Once Amended) The method of Claim 38, wherein the user-specified degree of
parallelism is specified in said statement.

93. (Once Amended) The method of Claim 38, wherein the user-specified degree of
parallelism is specified for operations that involve a particular table.

94. (Twice Amended) The computer-readable storage medium of Claim 74, wherein the user-
specified degree of parallelism is specified in said statement.

1 95. (Twice Amended) The computer-readable storage medium of Claim 74, wherein the user-
2 specified degree of parallelism is specified for operations that involve a particular table.

3